

AD-A149 818

ON THE LIMITING EMPIRICAL DISTRIBUTION FUNCTION OF THE  
EIGENVALUES OF A M. (U) PITTSBURGH UNIV PA CENTER FOR  
MULTIVARIATE ANALYSIS Z D BAI ET AL. OCT 84 TR-84-42  
AFOSR-TR-84-1215 F49620-82-K-0001

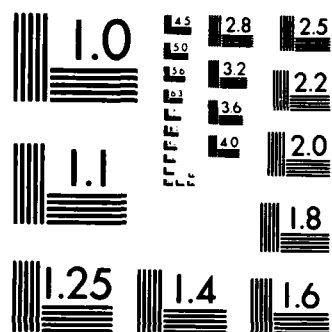
1/1

UNCLASSIFIED

F/G 12/1

NL





MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

3

AD-A149 818

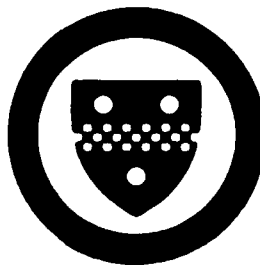
ON LIMITING EMPIRICAL DISTRIBUTION  
FUNCTION OF THE EIGENVALUES OF A  
MULTIVARIATE F MATRIX

Z. D. Bai<sup>1</sup>, Y. Q. Yin<sup>1</sup>, and  
P. R. Krishnaiah<sup>2</sup>

DTIC  
ELECTE  
JAN 29 1985  
S D

Center for Multivariate Analysis A

University of Pittsburgh



Approved for public release;  
distribution unlimited.

THIS FILE COPY

85 01 14 047

85 01 14 047

(3)

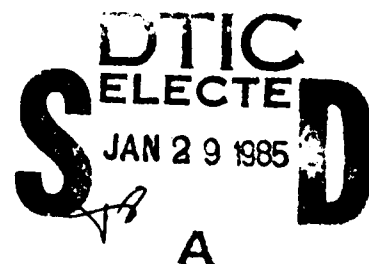
ON LIMITING EMPIRICAL DISTRIBUTION  
FUNCTION OF THE EIGENVALUES OF A  
MULTIVARIATE F MATRIX

Z. D. Bai<sup>1</sup>, Y. Q. Yin<sup>1</sup>, and  
P. R. Krishnaiah<sup>2</sup>

October 1984

Technical Report No. 84-42

Center for Multivariate Analysis  
Fifth Floor, Thackeray Hall  
University of Pittsburgh  
Pittsburgh, PA 15260



<sup>1</sup>Z. D. Bai and Y. Q. Yin are on a leave of absence from the China  
University of Science and Technology, Hefei, People's Republic of China

<sup>2</sup>The work of this author was supported by the Air Force Office of Scientific  
Research under Contract F49620-82-K-0001. Reproduction in whole or in part  
is permitted for any purpose of the United States Government.

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFOSR)  
NOTICE  
THIS  
...  
...  
...  
Chief, Technical ... Division

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

## REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION <b>UNCLASSIFIED</b>		1b. RESTRICTIVE MARKINGS													
2a. SECURITY CLASSIFICATION AUTHORITY		3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.													
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE		5. MONITORING ORGANIZATION REPORT NUMBER(S) <b>AFOSR-TR. 84-1215</b>													
4. PERFORMING ORGANIZATION REPORT NUMBER(S) <b>84-42</b> <b>18/</b>		7a. NAME OF MONITORING ORGANIZATION Air Force Office of Scientific Research													
6a. NAME OF PERFORMING ORGANIZATION University of Pittsburgh		7b. ADDRESS (City, State and ZIP Code) Directorate of Mathematical & Information Sciences, Bolling AFB DC 20332-6448													
6c. ADDRESS (City, State and ZIP Code) Center for Multivariate Analysis, Fifth Floor, Thackeray Hall, Pittsburgh PA 15260		9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F49620-82-K-0001													
8a. NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR		10. SOURCE OF FUNDING NOS. <table border="1"><tr><td>PROGRAM ELEMENT NO.</td><td>PROJECT NO.</td><td>TASK NO.</td><td>WORK UNIT NO.</td></tr><tr><td>61102F</td><td>2304</td><td>A5</td><td></td></tr></table>		PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.	61102F	2304	A5					
PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.												
61102F	2304	A5													
8b. OFFICE SYMBOL (If applicable) NM		11. TITLE (Include Security Classification) ON THE LIMITING EMPIRICAL DISTRIBUTION FUNCTION OF THE EIGENVALUES OF A MULTIVARIATE F MATRIX													
8c. ADDRESS (City, State and ZIP Code) Bolling AFB DC 20332-6448		12. PERSONAL AUTHOR(S) Z.D. Bai, Y.Q. Yin and P.R. Krishnaiah													
13a. TYPE OF REPORT Technical		13b. TIME COVERED FROM _____ TO _____													
14. DATE OF REPORT (Yr., Mo., Day) OCT 84		15. PAGE COUNT 13													
16. SUPPLEMENTARY NOTATION															
17. COSATI CODES <table border="1"><tr><td>FIELD</td><td>GROUP</td><td>SUB GR.</td></tr><tr><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td></tr></table>		FIELD	GROUP	SUB GR.										18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number) Canonical correlation analysis; empirical distribution function; large dimensional random matrices; limiting distribution; multivariate analysis of variance; (CONT.)	
FIELD	GROUP	SUB GR.													
19. ABSTRACT (Continue on reverse if necessary and identify by block number) In this paper, the authors derived an explicit expression for the limit of the empirical distribution function of a central multivariate F matrix when the number of variables and degrees of freedom tend to infinity in certain fashion. This distribution is useful in deriving the limiting distributions of certain test statistics which arise in multivariate analysis of variance, canonical correlation analysis and tests for the equality of two covariance matrices.  ITEM #18, SUBJECT TERMS, CONTINUED: multivariate F matrix.															
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED <input checked="" type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS <input type="checkbox"/>		21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED													
22a. NAME OF RESPONSIBLE INDIVIDUAL MAJ Brian W. Woodruff		22b. TELEPHONE NUMBER (Include Area Code) (202) 767- 5027													
		22c. OFFICE SYMBOL NM													

ON LIMITING EMPIRICAL DISTRIBUTION FUNCTION OF  
THE EIGENVALUES OF A MULTIVARIATE F MATRIX

Z. D. Bai, Y. Q. Yin and P. R. Krishnaiah

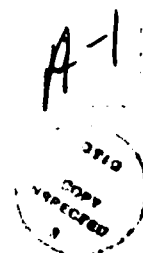
ABSTRACT

In this paper, the authors derived an explicit expression for the limit of the empirical distribution function of a central multivariate F matrix when the number of variables and degrees of freedom tend to infinity in certain fashion. This distribution is useful in deriving the limiting distributions of certain test statistics which arise in multivariate analysis of variance, canonical correlation analysis and tests for the equality of two covariance matrices.

**Keywords and Phrases:** Canonical correlation analysis, empirical distribution function, large dimensional random matrices, limiting distribution, multivariate analysis of variance, multivariate F matrix.

*large dimensional  
random matrices, short matrices, application  
to the F matrix*

Accession For	
PRA&I	<input checked="checked" type="checkbox"/>
TAB	<input type="checkbox"/>
Classification	
Distribution/	
Library Codes	
Author/Editor	
Subject	



## 1. INTRODUCTION

Various test procedures in multivariate analysis are based upon certain functions of the eigenvalues of random matrices. A considerable amount of work was done in the literature on the asymptotic distribution theory of these statistics when the sample size is very large. But, many situations arise in multivariate data analysis when the number of variables and the sample size are both very large. So, there is a great need to investigate the distributions of various functions of the eigenvalues of large dimensional random matrices. Distributions of the eigenvalues of large dimensional random matrices arise (e.g., see Mehta (1967)) in nuclear physics also.

Some work was done in the literature on the limiting empirical distribution function (e.d.f.) of large dimensional random matrices. Here, we note that the e.d.f. of a random matrix  $Z: p \times p$  is defined as  $N(x)/p$  where  $N(x)$  denotes the number of the eigenvalues of  $Z$  which are less than or equal to  $x$ . The e.d.f. (also known as spectral distribution) of  $Z$  is useful in deriving the distributions of certain functions of the eigenvalues of  $Z$ .

Now, let  $S_1: p \times p$  be distributed as central Wishart matrix with  $m$  degrees of freedom and  $E(S_1/m) = I_p$ . Also, let  $p$  and  $m$  both tend to infinity such that  $\lim(p/m) = y > 0$ . Then, it is known (see Grenander and Silverstein (1977), Jonsson (1982) and Wachter (1978)) that the e.d.f.  $F_m(x)$  of the eigenvalues of  $S_1/m$  tends to  $F_y(x)$  where  $F_y(x)$  is a distribution function with density function given by

$$f_y(x) = \begin{cases} \frac{\sqrt{(x-a)(b-x)}}{2\pi xy} & a < x < b \\ 0 & \text{otherwise} \end{cases}$$

where  $a = (1 - \sqrt{y})^2$ , and  $b = (1 + \sqrt{y})^2$  and  $0 < y \leq 1$ ; for  $1 < y < \infty$ ,  $F_y(x)$  has mass  $1 - (1/y)$  at zero and  $f_y(x)$  on  $(a, b)$ . Yin and Krishnaiah (1983b) showed that the spectral distribution of the sample covariance matrix has a limit when the underlying distribution is isotropic and  $y < 1$ .

Yin and Krishnaiah (1983a) showed that the spectral distribution of  $S_1 T / m$  tends to a limit in probability for each  $x$  under the following conditions:

- (a)  $T$  is a symmetric, positive definite matrix and  $G_p(x)$  is the e.d.f. of the eigenvalues of  $T$ ,
- (b)  $S_1$  and  $T$  are independent of each other,
- (c)  $\lim_{p, m \rightarrow \infty} (p/m) = y$  exists and finite
- (d)  $\int x^k dG_p(x) \rightarrow H_k$  exists in  $L^2$  for  $k = 1, 2, \dots$  and  $\sum H_{2k}^{-1/2} = \infty$ .

Yin and Krishnaiah (1984b) extended the above result to the case when  $S_1$  is the sample sums of squares and cross products matrix based upon observations from an isotropic population.

Now, let  $S_1: p \times p$  and  $S_2: p \times p$  be distributed independently as central Wishart matrices with  $m$  and  $n$  degrees of freedom and  $E(S_1/m) = E(S_2/n) = I_p$ . Then, the distribution of  $n S_1 S_2^{-1} / m$  is known to be the central multivariate  $F$  matrix. Applying the result of Yin and Krishnaiah (1983), Yin, Bai, and Krishnaiah (1983) showed that the limit of the spectral distribution of the central multivariate  $F$  matrix exists when  $p/m \rightarrow y'$  exists and  $(p/n) \rightarrow y < 1/2$  as  $p \rightarrow \infty$ . Silverstein (1984b) showed the validity of the above result even for the case  $1/2 \leq y < 1$  by making a minor modification in the proof of Yin, Bai, and Krishnaiah (1983). Yin and Krishnaiah (1983a) gave an expression for the moments of the limit of the e.d.f. of the eigenvalues of  $S_1 T / m$ . Starting from this expression, Silverstein (1984a) derived an explicit expression for the limit of the e.d.f. of the eigenvalues of the multivariate  $F$  matrix. In this paper, the authors give an alternative derivation of the above limit. The authors also gave explicit expressions for the moments of the above limiting distribution and these expressions are not known in the literature. Extensions of the results of this paper for nonnormal populations and nonnull cases are under investigation.



## 2. PRELIMINARIES

In this section, we give some results which are needed in the sequel as well as a brief review of known results on limiting spectral distribution of a multivariate F matrix.

**Lemma 2.1** Let  $z \in (0,1)$ ,  $a' = (1 - \sqrt{z})^2$  and  $b' = (1 + \sqrt{z})^2$ . If  $0 \leq |t| \leq a'$ , then

$$\frac{1}{2\pi z} \int_{a'}^{b'} \frac{1}{(x-t)} \{(x-a')(b'-x)\}^{\frac{1}{2}} dx = \frac{1}{2z} [1 + z - t - \{(1-z-t)^2 - 4tz\}^{\frac{1}{2}}] \quad (2.1)$$

**Proof** Making the transformation  $u = [2x - (b' + a')]/(b' - a')$  in the left side of (2.1), we obtain

$$R(t) = \frac{1}{2\pi z} \int_{a'}^{b'} \frac{1}{(x-t)} \{(x-a')(b'-x)\}^{\frac{1}{2}} dx = \frac{(b'-a')}{4\pi z} \int_{-1}^1 \frac{(1-u^2)^{\frac{1}{2}}}{u + \Delta} du \quad (2.2)$$

where  $\Delta = (b' + a' - 2t)/(b' - a')$ . It is known (see Jonsson (1982))

that  $R(0) = 1$ . So, for any  $r \in (0,1)$ , we have

$$\frac{1}{\pi\sqrt{r}} \int_{-1}^1 \frac{(1-u^2)^{\frac{1}{2}}}{u + ((1+r)/2\sqrt{r})} = 1. \quad (2.3)$$

Now, let  $\Delta = (1+r)/2\sqrt{r}$ . Since  $\Delta > 1$ , the condition  $r \in (0,1)$  is satisfied.

So, using (2.2) and (2.3), we obtain

$$R(t) = (b' - a')(\Delta - \sqrt{\Delta^2 - 1})/4z = [1 + z - t - \{(1-z-t)^2 - 4tz\}^{\frac{1}{2}}]/2z. \quad (2.4)$$

**Lemma 2.2** For any nonnegative integers  $m$  and  $w$ , we have

$$\sum_{\ell=0}^{[w/2]} (-1)^\ell \binom{m}{\ell} \binom{2m-2\ell}{2m-w} = 2^w \binom{m}{w}. \quad (2.5)$$

Here  $\binom{m}{w}$  is defined to be zero when  $m < w$ .

Proof If  $m = 0$ , the proof is trivial. We now prove the result for  $m > 0$  by induction. Suppose the result is true for a fixed value of  $m$ . Then, we have

$$\begin{aligned}
 \sum_{\ell=0}^{[w/2]} (-1)^\ell \binom{m+1}{\ell} \binom{2m+2-2\ell}{2m-w+2} &= \sum_{\ell=0}^{m+1} (-1)^\ell \left[ \binom{m}{\ell} + \binom{m}{\ell-1} \right] \binom{2m+2-2\ell}{2m+2-w} \\
 &= \sum_{\ell=0}^m (-1)^\ell \binom{m}{\ell} \left[ \binom{2m-2\ell}{2m-w} + 2 \binom{2m-2\ell}{2m+1-w} \right] \\
 &= 2^w \binom{m}{w} + 2 \cdot 2^{w-1} \binom{m}{w-1} = 2^w \binom{m+1}{w}.
 \end{aligned}$$

So the result follows.

### Lemma 3

$$\int \frac{(\alpha y + \beta) dy}{(\gamma + \delta y - \rho y^2)^{3/2}} = \frac{4\alpha\gamma - 2\beta\delta + (4\rho\beta + 2\alpha\delta)y}{(\delta^2 + 4\gamma\rho)\sqrt{\gamma + \delta y - \rho y^2}} + \text{const.} \quad (2.6)$$

The above lemma can be verified directly by differentiation.

### 3. MOMENTS OF THE LIMITING SPECTRAL DISTRIBUTION OF A MULTIVARIATE F MATRIX

Let  $F_{y,z}(x)$  denote the limit of the e.d.f. of the multivariate F matrix as defined in the preceding section. Also, let  $\{E_k\}_{k=1}^{\infty}$  denote the moments of  $F_{y,z}(x)$ . Then, from Yin and Krishnaiah (1983), we know that

$$E_k = \sum_{w=0}^{k-1} y^w \binom{k}{w} \frac{1}{(w+1)} B(k,w) \quad (3.1)$$

where

$$B(k,w) = \sum \frac{(w+1)!}{n_1! \dots n_{k-w}!} H_1^{n_1} \dots H_{k-w}^{n_{k-w}} \quad (3.2)$$

and the summation in (3.2) is over all possible values of  $n_1, \dots, n_{k-w}$  subject to the restrictions  $n_1 + \dots + n_{k-w} = w+1$  and  $n_1 + n_2 + \dots + (k-w)n_{k-w} = k$ . Also,  $H_i = E(x^{-i})$  for  $i = 1, 2, \dots$  where the density of  $x$  is given by

$$g_z(x) = \begin{cases} \frac{\sqrt{(x-a')(b'-x)}}{2\pi x z} & a' < x < b' \\ 0 & \text{otherwise} \end{cases} \quad (3.3)$$

where  $a'$ ,  $b'$ , and  $z$  are as defined in the preceding section. For any  $|t| < a'$ ,

$$\begin{aligned} \phi(t) &= tH_1 + t^2H_2 + t^3H_3 + \dots \\ &= E(t/(x-t)) \\ &= \frac{1}{2\pi z} \int_{a'}^{b'} [t\{(x-a')(b'-x)\}^{1/2}/(x-t)] dx \\ &= (1-z-t)[1 - \{1 - \{(4tz/(1-z-t)^2)\}^{1/2}\}]/2z \end{aligned} \quad (3.4)$$

by using Lemma 2.1. Since  $B(k,w)$  is the coefficient of  $t^k$  in Taylor expansion of  $\{\phi(t)\}^{w+1}$ , we obtain

$$\begin{aligned} B(k,w) &= \sum_{\ell=0}^{w+1} \binom{w+1}{\ell} (-1)^\ell \sum_{j=0}^k \binom{\ell}{j} (-1)^j 2^{j-w-1} z^{j-w-1} \\ &\quad \cdot \binom{w+1-2j}{k-j} (1-z)^{w+1-j-k} (-1)^{k-j} \end{aligned} \quad (3.5)$$

If  $w + 1 - 2j \geq 0$ ,  $\binom{w + 1 - 2j}{k - j} \neq 0$  implies  $0 \leq j \leq w + 1 - k \leq 0$ , i.e.  $j = 0$ .

Noticing  $\sum_{\ell=0}^{w+1} \binom{w+1}{\ell} (-1)^\ell = 0$ , we see that the expression of  $B(k, w)$  only contains

the terms with  $j > \frac{w+1}{2}$ . If  $\ell$  is even, then  $\binom{\ell}{j} \neq 0$  implies  $j \leq \frac{\ell}{2} \leq \frac{w+1}{2}$ , which is contrary to  $j > \frac{w+1}{2}$ . Thus, in the expression of  $B(k, w)$ , there are only the terms with  $j > \frac{w+1}{2}$  and  $\ell$  being odd. Applying lemma 2.2, we obtain

$$\begin{aligned}
 B(k, w) &= \sum_{\ell=0}^{[w/2]} \binom{w+1}{2\ell+1} (-1)^{k+1} \sum_{j=[\frac{w+3}{2}]^k}^k \binom{\ell + \frac{1}{2}}{j} 2^{2j-1-w} z^{j-w-1} \\
 &\quad \times \binom{w+1-2j}{k-j} (1-z)^{w+1-j-k} \\
 &= \sum_{\ell=0}^{[w/2]} \binom{w+1}{2\ell+1} \sum_{j=[\frac{w+3}{2}]^k}^k \frac{(2\ell+1)!(2j-2-2\ell)!}{j!\ell!(j-\ell-1)!} \\
 &\quad \times (-1)^{\ell} 2^{-w} z^{j-w-1} (1-z)^{w+1-j-k} \binom{k+j-w-2}{k-j} \\
 &= \sum_{j=w+1}^k \frac{(k+j-w-2)!(w+1)}{j!(j-w-1)!(k-j)!} z^{j-w-1} (1-z)^{w+1-j-k} \\
 &= \sum_{j=0}^{k-w-1} \frac{(k+j-1)!(w+1)}{(j+w+1)!j!(k-j-w-1)!} z^j (1-z)^{-k-j}
 \end{aligned} \tag{3.6}$$

From (3.5) and (3.6) it follows that

$$\begin{aligned}
 E_k &= \sum_{w=0}^{k-1} y^w \binom{k}{w} \sum_{j=0}^{k-w-1} \frac{(k+j-1)!}{j!(j+w+1)!(k-j-w-1)!} z^j (1-z)^{-k-j} \\
 &= \sum_{j=0}^{k-1} z^j (1-z)^{-k-j} \sum_{w=0}^{k-1-j} y^w \binom{k}{w} \frac{(k+j-1)!}{j!(j+w+1)!(k-j-w-1)!} .
 \end{aligned} \tag{3.7}$$

The  $k$ -th moment of  $\Gamma_{y,z}(\cdot)$  can be easily computed from (3.7)

#### 4. DERIVATION OF THE LIMITING SPECTRAL DISTRIBUTION OF MULTIVARIATE F MATRIX

Substituting the well-known formulas

$$\binom{k}{w} = \sum_{t=0}^w \binom{k-j-w-1}{t} \binom{j+w+1}{w-t} \quad (4.1)$$

into (3.7), we obtain the following by changing the order of summation:

$$\begin{aligned} E_k &= \sum_{j=0}^{k-1} z^j (1-z)^{-k-j} \sum_{t=0}^{\lfloor \frac{k-1-j}{2} \rfloor} \frac{(k+j-1)! y^t (1+y)^{k-j-1-2t}}{(k-j-1-2t)! t! (j+t+1)! j!} \\ &= \sum_{j=0}^{k-1} z^j (1-z)^{-k-j} \sum_{t=j}^{\lfloor \frac{k-j-1}{2} \rfloor} \frac{(k+j-1)! y^{t-j} (1+y)^{k+j-1-2t}}{(k+j-1-2t)! (t-j)! j! (t+1)!} \\ &= \sum_{t=0}^{k-1} \binom{2t}{t} \frac{1}{t+1} \sum_{j=0}^t z^j (1-z)^{-k-j} \binom{t}{j} \binom{k+j-1}{2t} y^{t-j} (1+y)^{k+j-1-2t} \end{aligned} \quad (4.2)$$

Using the formulas

$$\binom{k+j-1}{2t} = \sum_{s=0}^{2t} \binom{k-1}{s} \binom{j}{2t-s} \quad (4.3)$$

in (4.2), we obtain

$$\begin{aligned} E_k &= \sum_{t=0}^{k-1} \binom{2t}{t} \frac{1}{t+1} \sum_{s=t}^{2t} \binom{k-1}{s} \sum_{j=0}^t z^j (1-z)^{-k-j} \\ &\quad \times y^{t-j} (1+y)^{k+j-1-2t} \binom{s-t}{t-j} \binom{t}{s-t} \\ &= \sum_{t=0}^{k-1} \binom{2t}{t} \frac{1}{t+1} \sum_{s=t}^{2t} \binom{k-1}{s} \binom{t}{s-t} \sum_{j=0}^{s-t} \binom{s-t}{j} z^{t-j} (1-z)^{-k-t+j} \\ &\quad \times y^j (1+y)^{k-t-1-j} \\ &= \sum_{t=0}^{k-1} \binom{2t}{t} \frac{1}{t+1} \sum_{s=t}^{2t} \binom{k-1}{s} \binom{t}{s-t} z^{2t-s} (1-z)^{-k-t} \\ &\quad \times (1+y)^{k-s-1} (y+z)^{s-t} \end{aligned}$$

$$= \sum_{t=0}^{k-1} \binom{2t}{t} \frac{1}{t+1} \sum_{s=0}^t \binom{k-1}{s+t} \binom{t}{s} z^{t-s} (1-z)^{-k-t} \\ \times (1+y)^{k-t-s-1} (y+z)^s \quad (4.4)$$

$$= \sum_{s=0}^{[(k-1)/2]} \sum_{t=s}^{k-1-s} \binom{k-1}{s+t} \binom{t}{s} \binom{2t}{t} \frac{1}{t+1} z^{t-s} (1-z)^{-k-t} \\ \times (1+y)^{k-s-1-t} (y+z)^s \\ = \sum_{s=0}^{[(k-1)/2]} \sum_{t=0}^{k-1-2s} \binom{k-1}{2s} \binom{k-1-2s}{t} \frac{(2s)!(2s+2t)!}{s!(2s+t)!(s+t+1)!} \\ \times z^t (1-z)^{-k-t-s} (1+y)^{k-2s-1-t} (y+z)^s \quad (4.5)$$

Define a random vector  $(U, V)$  where the marginal density of  $U$  is  $\frac{2}{\pi} (1-x^2)^{-1/2} I_{(0,1)}(x)$  and the conditional density of  $V$  given  $U = x$  is

$$\frac{|v|}{1-x^2} I_{[-\sqrt{1-x^2}, \sqrt{1-x^2}]}(v). \quad (4.6)$$

It is easy to see that

$$E\{U^{2s} V^{2t+2s}\} = \frac{(2s)!(2t+2s)!}{s!(2s+t)!(s+t+1)!} 4^{-2s-t} \quad (4.7)$$

and that

$$E\{U^{2s+1} V^{2t+2s+1}\} = 0.$$

Hence, from (4.5), (4.7) it follows that

$$E_k = E \sum_{s=0}^{k-1} \sum_{t=0}^{k-1-s} \binom{k-1}{s} \binom{k-1-s}{t} z^t (1-z)^{-k-t-\frac{s}{2}} \\ \times (1+y)^{k-s-t-1} (y+z)^{\frac{s}{2}} U^s V^{2t+s} \\ = \frac{1}{(1-z)^k} E \left[ 1+y + \frac{4V^2 z}{1-z} + \frac{4UV\sqrt{y+z}}{\sqrt{1-z}} \right]^{k-1} \quad (4.8)$$

Now, we compute the distribution of

$$\left[ 1 + y + \frac{4V^2 z}{1-z} + \frac{4UV \sqrt{y+3}}{\sqrt{1-z}} \right] / (1-z). \quad (4.9)$$

Let  $w_1 = UV$  and  $w_2 = V^2$ . Then we can easily show that the joint distribution of  $(w_1, w_2)$  is

$$\frac{1}{\pi} w_2 (w_2 - w_1^2)^{-3/2} I_{[w_1^2 + w_2^2 \leq w_2]}$$

where  $I_{[\alpha \leq \beta]}$  takes value 1 if  $\alpha \leq \beta$  is true and zero otherwise. Applying Lemma 2.3, we can compute the density of  $\frac{4\sqrt{y+z}}{\sqrt{1-z}} w_1 + \frac{4z}{1-z} w_2$  as follows:

Let  $q_1 < q_2$  be the two roots of the equation

$$\frac{(1-z)^2}{16z^2} q^2 + \frac{1-z}{16(y+z)} (x-q)^2 = \frac{1-z}{4z} q \quad (4.10)$$

in the variable  $q$ . Let  $\alpha = 1$ ,  $\beta = 0$ ,  $r = \{(1-z)/16(y+z)\}x^2$ ,  $\delta = \frac{1-z}{4z} + \frac{1-z}{8(y+z)}x$  and  $\rho = \frac{1-z}{16(y+z)}$ . Then

$$\Delta_1 = 4\alpha r - 2\beta\delta = -\frac{1-z}{4(y+z)} x^2,$$

$$\Delta_2 = 4\beta\rho + 2\alpha\delta = \frac{1-z}{2z} + \frac{1-z}{4(y+z)} x$$

$$\Delta_3 = \delta^2 + 4\gamma\rho = \frac{(1-z)^2[y+z+xz]}{16z^2(y+z)}.$$

By Lemma 2.3, we get the density of  $\frac{4\sqrt{y+z}}{\sqrt{1-z}} w_1 + \frac{4z}{1-z} w_2$  as given below:\*

$$\begin{aligned} f_1(x) &= \frac{1}{\pi} \frac{\sqrt{1-z}}{4\sqrt{y+z}} \frac{(1-z)^2}{16z^2} \int_{q_1}^{q_2} q \left( \frac{1-z}{4z} q - \frac{1-z}{16(y+z)} (x-q)^2 \right)^{-3/2} dq^{**} \\ &= \frac{(1-z)^{5/2}}{64\pi z^2 \sqrt{y+z}} \frac{\Delta_1 + \Delta_2 q}{\Delta_3 \sqrt{\frac{(1-z)}{4z} q - \frac{1-z}{16(y+z)} (x-q)^2}} \Big|_{q_1}^{q_2}. \quad (4.11) \end{aligned}$$

Since  $q_1$  and  $q_2$  are roots of equation (4.10), we have

\* Use the formula  $f_{x+y}(x) = \int f(x-q, q) dq$  where  $f_{x+y}(\cdot)$  and  $f(\cdot, \cdot)$  are the densities of  $x+y$  and  $(x, y)$  respectively.

\*\* Note that the integrand is zero outside the interval  $[q_1, q_2]$  by the indicator factor.

$$\begin{aligned}
 f_1(x) &= \frac{(1-z)^{5/2}}{64\pi z^2 \sqrt{y+z}} \frac{\Delta_1 + \Delta_2 q}{\Delta_3 \frac{1-z}{4z} q} \bigg|_{q_1}^{q_2} * \\
 &= \frac{(1-z)^{5/2}}{64\pi z^2 \sqrt{y+z}} \frac{(-\Delta_1)}{\Delta_3 \frac{1-z}{4z}} \frac{q_2 - q_1}{q_1 q_2}. \quad (4.12)
 \end{aligned}$$

From (4.10), we can compute

$$q_1 q_2 = \frac{z^2 x^2}{y+z-yz} \quad (4.13)$$

$$q_2 - q_1 = \frac{2z \sqrt{(y+z)(-x^2(1-z) + 4xz + 4(y+z))}}{y+z-yz} \quad (4.14)$$

From (4.12) - (4.14), it follows that

$$\begin{aligned}
 f_1(x) &= \frac{(1-z)^{5/2}}{64\pi z^2 \sqrt{y+z}} \frac{\left[ x^2(1-z)/4(y+z) \right] 2z \sqrt{(y+z)(-x^2(1-z) + 4xz + 4(y+z))}}{(1-z)^2 [y+z+xz] \frac{1-z}{4z} z^2 x^2} \\
 &= \frac{\sqrt{1-z}}{2\pi(y+z+xz)} \sqrt{-x^2(1-z) + 4xz + 4(y+z)}. \quad (4.15)
 \end{aligned}$$

From this we can easily obtain the density of  $\left[ 1+y + \frac{4V^2 z}{1-z} + \frac{4UV\sqrt{y+z}}{\sqrt{1-z}} \right] / (1-z)$  as

$$\begin{aligned}
 f_2(x) &= \frac{(1-z) \sqrt{-(1-z)^2 x^2 + 2(1+y+z-yz)x - (1-y)^2}}{2\pi(xz+y)} \\
 &= \frac{(1-z)^2 \sqrt{(x-a)(b-x)}}{2\pi(xz+y)} \quad (4.16)
 \end{aligned}$$

$$\text{where } a = \frac{(1-\sqrt{y+z-yz})^2}{(1-z)^2} \text{ and } b = \frac{(1+\sqrt{y+z-yz})^2}{(1-z)^2}.$$

Since  $f_1(x) \neq 0$  iff equation (4.10) has two different roots, we find by checking the steps of computation that  $f_2(x) \neq 0$  iff  $a < x < b$ . Recalling (4.9), we obtain

$$\begin{aligned}
 E_k &= \frac{1}{1-z} \int_a^b x^{k-1} f_2(x) dx \\
 &= \int_a^b x^k \left[ \frac{f_2(x)}{x(1-z)} \right] dx. \quad (4.17)
 \end{aligned}$$

\* Note that  $\frac{(1-z)}{4z} q - \frac{1-z}{16(y+z)} (x-q)^2 = \left( \frac{1-z}{4z} q \right)^2$  for  $q = q_1$  or  $q_2$ .



Now, let

$$f(x) = \frac{f_2(x)}{x(1-z)} = \begin{cases} \frac{(1-z)\sqrt{(x-a)(b-x)}}{2\pi x(y+xz)} & \text{if } a < x < b, \\ 0 & \text{otherwise.} \end{cases} \quad (4.18)$$

In [7], it is shown that the distribution  $F_{y,z}$  is determined by all its moments.

From (4.11), (4.12) it follows that

$$F_{y,z}(x) = \gamma I_{\{(0,\infty)\}}(x) + \int_{-\infty}^x f(x) dx. \quad (4.19)$$

where  $\gamma = 1 - \int_{-\infty}^{\infty} f(x) dx$ .

Finally, we only need to compute the integral  $\int_a^b f(x) dx$ . If we set

$U = [2x - (b+a)]/(b-a)$ , we get

$$\begin{aligned} I &= \int_a^b f(x) dx = \frac{(1-z)}{2\pi y} \left[ \int_a^b \frac{\sqrt{(x-a)(b-x)}}{x} dx - \int_a^b \frac{\sqrt{(x-a)(b-x)}}{x + \frac{y}{z}} dx \right] \\ &= \frac{(1-z)(b-a)}{4\pi y} \left[ \int_{-1}^1 \frac{\sqrt{1-u^2}}{u + \lambda_1} du - \int_a^b \frac{\sqrt{1-u^2}}{u + \lambda_2} du \right], \end{aligned}$$

where  $\lambda_1 = \frac{b+a}{b-a}$  and  $\lambda_2 = \frac{b+a+(2y)/z}{b-a}$ . Using (2.2) we get

$$\begin{aligned} I &= \frac{(1-z)(b-a)}{4y} \left[ (\lambda_1 - \sqrt{\lambda_1^2 - 1}) - (\lambda_2 - \sqrt{\lambda_2^2 - 1}) \right] \\ &= \frac{1-z}{2y} \left[ \sqrt{ab + \frac{y}{z}(a+b) + \frac{b^2}{z^2}} - \sqrt{ab} - \frac{y}{z} \right] \\ &= \frac{1-z}{2y} \left[ \frac{y+z}{z(1-z)} - \frac{|1-y|}{1-z} - \frac{y}{z} \right] \\ &= \frac{1}{2y} [1+y - |1-y|] \\ &= \begin{cases} 1 & \text{if } 0 < y \leq 1 \\ \frac{1}{y} & \text{if } y > 1. \end{cases} \end{aligned}$$

Hence

$$\gamma = \begin{cases} 0 \\ 1 - \frac{1}{y} \end{cases}$$

Substituting this into (4.19), we get the expression of  $F_{y,z}(x)$ .

## ADDED IN PROOF

After this paper was typed, the following interesting paper was brought to the attention of the authors:

Wachter, K. W. (1980). The limiting empirical measure of multiple discriminant ratios. Annals of Statistics, 8, 937-957.

In the above paper, Wachter gave an explicit expression for the e.d.f. of  $(XJX')(XX')^{-1}$  in the limiting case when  $p \rightarrow \infty$ ,  $J$  is a projection matrix and the columns of  $X$ :  $p \times (n+m)$  are distributed independently as multivariate normal with mean vector  $0$  and covariance matrix  $I$ . When

$$J = \begin{pmatrix} I_m & 0 \\ 0 & 0 \end{pmatrix},$$

we observe that  $(XJX')(XX')^{-1} = (X_1X_1')(X_1X_1' + X_2X_2')^{-1}$  where  $X = (X_1 \ X_2)$

The random matrix considered by us is equivalent to  $(X_1X_1')(X_2X_2')^{-1}$  and so our result on the limiting e.d.f. can be obtained from the result of Wachter.

But, the method used by us is entirely different from the method used by Wachter (1980). Also, we derived the moments of the above distribution and these expressions are new.

## REFERENCES

- [1] GRENANDER, U. and SILVERSTEIN, J. W. (1977). "Spectral analysis of networks with random topologies", SIAM J. Appl. Math., 32, 499-519.
- [2] JONSSON, D. (1982). "Some limit theorems for the eigenvalues of a sample covariance matrix", J. Multivariate Anal., 12, 1-38.
- [3] MEHTA, M. L. (1967). Random Matrices and the Statistical Theory of Energy Levels. Academic Press, Inc.
- [4] SILVERSTEIN, J. W. (1984a). "The limiting eigenvalue distribution of a multivariate F matrix", To appear in SIAM J. Math. Anal.
- [5] SILVERSTEIN, J. W. (1984b). "Comments on a result of Yin, Bai and Krishnaiah", J. Multivariate Anal., In press.
- [6] WACHTER, K. W. (1978). "The strong limits of random spectra for sample matrices of independent elements", Ann. Prob., 6, 1-18.
- [7] YIN, Y. Q., BAI, Z. D. and KRISHNAIAH, P. R. (1983). "Limiting behavior of the eigenvalues of a multivariate F matrix", J. Multivariate Anal., 13, 508-516.
- [8] YIN, Y. Q. and KRISHNAIAH, P. R. (1983a). "A limit theorem for the eigenvalues of product of two random matrices. J. Multivariate Anal., 13, 489-507.
- [9] YIN, Y. Q. and KRISHNAIAH, P. R. (1983b). "Limit theorems for the eigenvalues of the sample covariance matrix when the underlying distribution is isotropic", To appear in Teoriya Veroyatnostei i ee Primeneniya.
- [10] YIN, Y. Q. and KRISHNAIAH, P. R. (1983c). "Limit theorems for the eigenvalues of product of large dimensional random matrices when the underlying distribution is isotropic", To appear in Teoriya Veroyatnostei i ee Primeneniya.

**END**

**FILMED**

**3-85**

**DTIC**